

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
12 April 2001 (12.04.2001)

PCT

(10) International Publication Number  
**WO 01/25485 A2**

(51) International Patent Classification<sup>7</sup>: C12Q 1/68

(21) International Application Number: PCT/US00/27319

(22) International Filing Date: 4 October 2000 (04.10.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/157,937 6 October 1999 (06.10.1999) US

(71) Applicant: AMERSHAM PHARMACIA BIOTECH,  
INC. [US/US]; 800 Centennial Avenue, Piscataway, NJ  
08855-1327 (US).

(71) Applicant and

(72) Inventor: ZHAO, Xiaodong [CN/US]; 37 Shaffer Road,  
Bridgewater, NJ 08807 (US).

(74) Agent: RONNING, Royal, N., Jr.; Amersham Pharma-  
cia Biotech, Inc., 800 Centennial Avenue, Piscataway, NJ  
08855-1327 (US).

(81) Designated States (national): AE, AL, AM, AT, AU, AZ,  
BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK,  
DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL,  
IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU,  
LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT,  
RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA,  
UG, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM,  
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian  
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European  
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,  
IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG,  
CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— Without international search report and to be republished  
upon receipt of that report.

For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.



(54) Title: METHOD FOR DETECTING MUTATIONS USING ARRAYED PRIMER EXTENSION

**WO 01/25485 A2**

(57) Abstract: A method for detecting a mutation in a target nucleic acid sequence that comprises attaching oligonucleotide primers to a substrate, wherein the oligonucleotide primers have a sequence that is complementary to the target nucleic acid sequence, and wherein the oligonucleotide primers are grouped according to the identity of the first base which would be expected to be added to the primer through the process of primer extension; hybridizing to the oligonucleotide primers a sample nucleic acid sequence which possibly contains a mutation; extending each oligonucleotide primer by one base using a reaction mixture comprising labeled ddNTPs; and detecting a mutation in the sample nucleic acid sequence by detecting the presence of a labeled ddNTP which does not correspond to the identity of the base expected to be added to the primer through the process of primer extension. Mutations may be detected in mRNA or DNA. The labeled ddNTPs may be labeled with a fluorescent dye, a chemiluminescent reagent, a radioactive label, or an electrically conductive tag. A nucleic acid array and a kit for detecting genetic mutations are also disclosed. The instant disclosure also pertains to a method for detecting a mutation in a target nucleic acid sequence that comprises identifying the base expected to be added to a primer located at a particular coordinate on an oligonucleotide array as expected from the target nucleic acid sequence; identifying the base actually added to the primer located at the particular coordinate on the oligonucleotide array through the process of primer extension; comparing the base actually added to the primer at the particular coordinate with the base expected to be added to the primer at the particular coordinate; and reporting those instances where the bases are not the same, in order to identify a mutation.

METHOD FOR DETECTING MUTATIONS USING ARRAYED PRIMER  
EXTENSION

BACKGROUND OF THE INVENTION

Upon completion of the human genome project, genotyping of individuals will become possible and important to trace biological reactions of each individual. This is one of the conceptions for health care in the future. Molecular medicine is expected to evolve for each individual based on his or her genetic identity. Chip-based nucleic acid sequence analysis is a powerful tool for re-sequencing gene to identify single nucleotide polymorphisms (SNP), which is believed to be responsible for individual identity. Spatially addressed primer array offers a cheap and effective method for SNP detection. However, to justify the signal of each spot in the entire array and to analyze the complete data is very tedious and requires decent software. Herein, I describe a novel format of arrayed primer extension, which simplifies the data analysis and leads to straightforward answer.

The current art of spotting microarray on a chip has the oligonucleotides laid on a chip sequentially from one end (3' or 5') to the other end (5' or 3') of the target. As shown in Figure 1, the primers add one base difference as they proceed from one end to the other end. A complete hybridization map to the arrayed primers reveals the sequence of the gene in question.

In a conventional oligonucleotide array (for hybridization with a nucleic acid and followed by a primer extension reaction), each of the four terminators are labeled with different fluorescent dyes. Thus, each terminator that has been incorporated onto the oligonucleotide primer may be identified on the image by color. However, it is often difficult to ascertain the presence of mutations in samples using such sequentially placed oligonucleotide primer arrays without the aid of complex software.

While nucleotide arrays present many advantages in the analysis of nucleic acid sequences, improvements in the techniques incorporating such arrays are still possible.

## BRIEF SUMMARY OF THE INVENTION

A method for detecting a mutation in a target nucleic acid sequence that comprises: attaching oligonucleotide primers to a substrate, wherein the oligonucleotide primers have a sequence that is complementary to the target nucleic acid sequence, and wherein the oligonucleotide primers are grouped according to the identity of the first base which would be expected to be added to the primer through the process of primer extension; hybridizing to the oligonucleotide primers a sample nucleic acid sequence which possibly contains a mutation; extending each oligonucleotide primer by one base using a reaction mixture comprising labeled ddNTPs and enzyme; and detecting a mutation in the sample nucleic acid sequence by detecting the presence of a labeled ddNTP which does not correspond to the identity of the base expected to be added to the primer through the process of primer extension. The sample nucleic acid sequence may be mRNA or DNA. The labeled ddNTPs may be labeled with a fluorescent dye, a chemiluminescent reagent, a radioactive label, a redox tag, or an electrically conductive tag.

The instant invention also pertains to a nucleic acid array comprising a substrate containing oligonucleotide primers having a sequence complementary to a target nucleic acid sequence, wherein the oligonucleotide primers are grouped according to the identity of the first base which would be expected to be added to the primer through the process of primer extension after hybridizing with a sample nucleic acid sequence.

The instant invention also pertains to a kit for detecting genetic mutations comprising the nucleic acid array as described above.

The instant invention also pertains to a method for detecting a mutation in a target nucleic acid sequence that comprises: identifying the base expected to be added to a primer located at a particular coordinate on an oligonucleotide array as expected from the target nucleic acid sequence; identifying the base actually added to the primer located at the particular coordinate on the oligonucleotide array through the process of primer extension; comparing the base actually added to the primer at the particular coordinate with the base expected to be added to the primer at the particular coordinate; and

reporting those instances where the bases are not the same, in order to identify a mutation.

#### BRIEF DESCRIPTION OF THE DRAWINGS

FIGURE 1 shows a conventionally arranged oligonucleotide array.

FIGURE 2 shows one possible reaction scheme for preparing a bromoacetamide derivatized silane glass surface.

FIGURE 3a shows one embodiment of the oligonucleotide primer as a tail structure having 5 groups capable of reacting with the surface of the substrate.

FIGURE 3b shows one embodiment of the oligonucleotide primer having a closed loop structure.

FIGURE 3c shows one embodiment of the oligonucleotide primer as a hairpin molecule containing six nucleosides connected by thiophosphates in the loop and having a single-stranded 3' end and a 5' end in the double-stranded stem.

FIGURE 3d shows a detailed view of the oligonucleotide primer depicted in FIGURE 3c.

FIGURE 4 shows the attachment reaction between the oligonucleotide primer depicted in Figure 3 with the bromoacetamide derivatized silane glass surface shown in Figure 2.

FIGURE 5 shows one embodiment of an arrayed primer extension reaction involving certain oligonucleotide primers described herein.

FIGURE 6 shows one embodiment of a ligation reaction involving certain oligonucleotide primers described herein.

## DETAILED DESCRIPTION OF THE INVENTION

Proposed herein is an array of primers that segregated by the base to be incorporated in the extension reaction. The entire array has four blocks, one for each of the incorporated bases, A, C, G, and T in the order in which the bases appear in the target gene.

Block A has sequentially addressed primers that are terminated with A, block C has sequentially addressed primers terminated with C. It is the same with block G and T. Therefore one can assemble the gene by base picking from the four blocks. If a mutation (or SNP) in a gene exists, it is revealed as a missing or mixed color spot in one block and confirmed as a missing or mixed color spot in the complementary base-pair block. This new format for arrayed primer extension would simplify the software requirement for deconvolution data analysis and visual recognition.

A particularly preferred method for anchoring oligonucleotides to the substrate is taught in U.S. application serial number 09/388,702 filed on September 2, 1999, entitled "Method for Anchoring Oligonucleotides to a Substrate," the entire disclosure of which is incorporated herein by reference. The substrate may be selected from a variety of materials, including glass, inorganic or organic polymer, and metal. Preferably, the substrate is glass. Although any type of glass may be used as a substrate, the preferred substrate is borosilicate glass. The substrate may take various physical forms, such as (but not limited to) slides or beads.

The surface of the substrate is modified to facilitate attachment of the oligonucleotide to the surface. The oligonucleotide is generally modified with a nucleophile, such as aminoalkyl, sulfhydryl and thiophosphate groups. The surface of the solid support is modified with a corresponding reactive group, such as substituted benzenesulfonate, methanesulfonate, and trifluoromethanesulfonate as leaving groups for the nucleophile on the primer. The reactive groups on the solid support can also be maleimide, vinylsulfonate and acrylate as electrophilic C=C double bonds for a nucleophile, and disulfide modification for disulfide exchange. The oligonucleotide can also be modified with biotin and the surface of the substrate coated with streptavidin or avidin. Preferred reactive groups for the oligonucleotide include thiophosphate, aminoalkyl and sulfhydryl. Most preferably, the reactive groups are thiophosphate.

Multiple reactive groups on the oligonucleotide have been found to enhance reaction efficiency. Preferably, the subject oligonucleotides contain multiple groups in a single molecule capable of reacting with the surface of the substrate. In general, the oligonucleotide can contain any number of reactive groups. Preferably, the oligonucleotide contains from one to five reactive groups (see Figure 3a which shows 5 reactive groups). The reactive groups may be arranged within one nucleoside of one another, or spaced throughout the oligonucleotide. Preferably, the reactive groups are arranged within two to six nucleosides of one another to create a hairpin structure in the oligonucleotide. The reactive groups may be the same or different. Although the subject oligonucleotides are not limited to those oligonucleotides having hairpin configurations; oligonucleotides having hair pin configurations such as those depicted in Figures 3c and 3d are preferred. Alternatively, the subject oligonucleotide may form a closed loop structure as shown in Figure 3b.

In a preferred embodiment, the glass substrate contains a bromoacetamide derivatized silane glass surface. Figure 2 depicts one possible reaction scheme for preparing a bromoacetamide derivatized silane glass surface. The reactive slides can be prepared by starting from either regular glass slides or amine coated slides (commercially available from Amersham Pharmacia Biotech Inc and Corning Inc). The two-phase reaction is very efficient.

One preferred embodiment of the oligonucleotide is shown in Figure 3c, and in greater detail in Figure 3d. In Figures 3c and 3d, the oligonucleotide primer is a hairpin molecule containing six nucleosides in the loop. The nucleosides in the loop are connected by thiophosphates. The hairpin has a single-stranded 3' end and a 5' end in the double-stranded stem. The single-stranded tail is equivalent to a single-stranded oligonucleotide. This approach allows fast and efficient attachment by maintaining relatively low concentrations of oligonucleotide primers and high concentrations of reactive groups. Attachment of the primer to the bromoacetamide derivatized silane glass surface is shown in Figure 4.

Because the oligonucleotide can attach to the substrate at any one of the reactive sites in the loop, the reaction probability (reaction rate and efficiency) increases with the number of reactive sites contained in the loop. In addition, as the anchorage of the

oligonucleotide is in the backbone, both 5' and 3' ends are free for modifications with reporter groups. This is more versatile than when the anchors are at either 5' or 3' end. While a reaction occurs at one end, the other end is available for further labeling. Moreover, the hairpin structure has a stem-loop moiety, which adds extra parameters to space the oligonucleotides from each other and control the density and desired upward conformation of the single-stranded moiety for primer extension as well as ligation.

This attachment scheme offers significant advantages for immobilization of cDNA or any PCR product to microchips by hybridization and ligation (see Figures 4 and 5). For example, if the 5' end of the hairpin is phosphorylated chemically during synthesis and the 3' single-stranded tail has a sequence which is complementary to a PCR primer sequence, the PCR product could hybridize with the single-stranded tail and be linked covalently by enzymatic or chemical ligation. Many of the current approaches for covalently attaching a PCR product to a surface focus mainly on modifying PCR primers. One disadvantage of this approach is that the reactive position may be buried in the self-folded macromolecule. The approach of hybridization followed by ligation may have a higher reaction probability than the simple interaction of reactive groups, since such bimolecular reaction depends on more factors than the sole collision probability of two substrates. The instant method and oligonucleotides are not limited to the hybridization of nucleic acids. For example, the oligonucleotide primer may also contain a nucleotide sequence corresponding to a protein binding site and would be useful, for example, in protein assays.

The instant attachment method has the potential to immobilize oligonucleotides and DNAs by a covalent bond to a silica based glass surface with improved control of reaction efficiency, molecular density and conformation, and enzyme compatibility. Given the versatility of the instant oligonucleotides, it is possible that uses may be found with substrates other than those mentioned.

The instantly configured arrays may be used to detect mutations in sample nucleic acid sequences such as mRNA or DNA.

Terminators which are useful in the practice of the instant method may be selected from the group consisting of deoxynucleoside triphosphates and dideoxynucleoside triphosphates. Naturally occurring terminators, such as ddNTPs, may

be used; however, synthetically-modified terminators, such as those containing a further methyl group blocking the 3' hydroxyl may also be used. The terminators useful in the practice of the instant method may be labeled in a variety of ways. For example, the labeled terminators may be labeled with a fluorescent dye, a chemiluminescent reagent, or a radioactive label. The terminators may also be labeled with a redox tag. For example, an organometallic tag, such as ferrocene, may be attached to the terminator and the change in electric potential at a particular spot on the microarray may be measured using a microelectrode. The labeled terminators may also be labeled with an electrically conductive tag and the change in conductivity at a particular spot on the microarray may be measured using a microelectrode.

Variations of the instant method may be used with conventionally configured oligonucleotide arrays as well, i.e., those in which the oligonucleotides are laid on a chip sequentially from one end (3' or 5') to the other end (5' or 3') of the target, and one base difference is added as the oligonucleotides are arranged from one end of the chip to the other end (e.g., the oligonucleotide array configuration depicted in Figure 1). Because the target nucleic acid sequence is already known, we are able to predict the identity of the base which is expected to be added to a primer that is located at a particular coordinate on an oligonucleotide array. This information may be entered into a computer program. After hybridization and subsequent primer extension, it is possible to identify the base actually added to the primer located at the particular coordinate on the oligonucleotide array. This information may also be entered into the computer program. If we compare the base actually added to the primer at the particular coordinate with the base that was expected to be added to the primer at the particular coordinate (e.g., using the computer program), we will be able to detect mutations in a target nucleic acid sequence by identifying those instances where the base actually added to the primer is not the same as the base which was expected to be added to the primer at that particular coordinate on the array.

The instantly disclosed method for detecting mutations using arrayed primer extension has simplicity of experiment, software development and data analysis. This method allows the researcher to immediately identify mutations using the naked eye. It can eliminate de-convolution errors by calibrating multiple standards and serve to reduce

overall error rate for application in diagnostics. This conception could also be used for software design even for an array prepared as the conventional format. The following examples are for illustration purposes only and should not be used in any way to limit the appended claims.

## EXAMPLES

### Slide Preparation

Pre-washed glass slides (25x75 mm microscope slides from VWR Scientific Products West Chester, PA) were immersed for 3 minutes in 95% ethanol (350 mL) containing approximately 2% (wt/vol) 3-aminopropylmethylethoxysilane, pH 5.0, washed with ethanol, and cured in an oven at 75 °C for 4 hours. After curing, the glass slides were placed on a glass rack and immersed in 160 mL of N, N-dimethylformamide (DMF) containing bromoacetic acid (0.45 g, 3.3 mmol), 4-(dimethylamino)-pyridine (DMAP) (0.04 g, 0.3 mmol), and 1,3-dicyclohexylcarbodiimide (DCC) (0.66 g, 3.2 mmol). The reaction was stirred in the dark for 2 hours. The slides were then washed with ethanol and air-dried.

### Array Preparation

Four primers were specifically designed for sense strand of human p53 gene exon 6. The primers were purchased from Sigma-Genosys (The Woodland ,Texas). Their sequences are as below:

- (1) aaaTGGCCCCTCCTCAGCATCTTATCCG
- (2) aaaTCACTGATTGCTCTTAGGTCTGGCC
- (3) aaaTTCCTCACTGATTGCTCTTAGGTCT
- (4) aaaACTGATTGCTCTTAGGTCTGGCCCC

Three adjacent thiophosphates are located to the right of lower case a's in each primer sequence. The hairpin oligonucleotide was dissolved and stored in 1×TE (10 mM Tris-HCl, 1 mM EDTA, pH 7.0). The concentration of oligonucleotide was calculated based on spectrophotometric measurement. The primers (50 µM) were spotted on a bromoacetamide silane coated glass surface using a Molecular Dynamics Gen III spotter (Molecular Dynamics, Sunnyvale, CA). The volume of the spots was 0.7 nL and the size

was 130  $\mu\text{m}$  in diameter. The primers were spotted in four blocks on a chip. Primer (1) was spotted in an 8x8 spot block, which can be extended by a DNA polymerase with dye labeled ddATP. Primer (2) was spotted in another block, which can be extended by a DNA polymerase with dye labeled ddCTP. Similarly, primers (3) and (4) were spotted in a third and fourth block respectively, which can be extended by a DNA polymerase with dye labeled ddGTP and ddTTP respectively.

#### Primer extension

The slide with primers was placed on a hot plate heated to 48 °C. Primer extension was carried out in 45  $\mu\text{L}$  of Thermo Sequenase buffer (25 mM TAPS buffer, pH 9.3, 50 mM KCl, 2 mM MgCl<sub>2</sub>, 1mM 2-mercaptoethanol) with 0.5  $\mu\text{M}$  of dye terminators Cy5-ddATP, Cy3-ddCTP, Cy5.5-ddGTP, and Cy3.5-ddTTP, and Thermo Sequenase DNA polymerase (4 unit). The templates were PCR products of human p53 gene exon 6. The reaction mixture was washed away with boiling water after being incubated for 10 minutes. The image of the dye terminator labeled spots was obtained using an Avalanche microscanner (Molecular Dynamics, Sunnyvale, CA), and exhibited resolution.

Although various embodiments of the instant invention are described in detail above, the instant invention is not limited to such specific examples. Various modifications will be readily apparent to one of ordinary skill in the art and fall within the spirit and scope of the following appended claims.

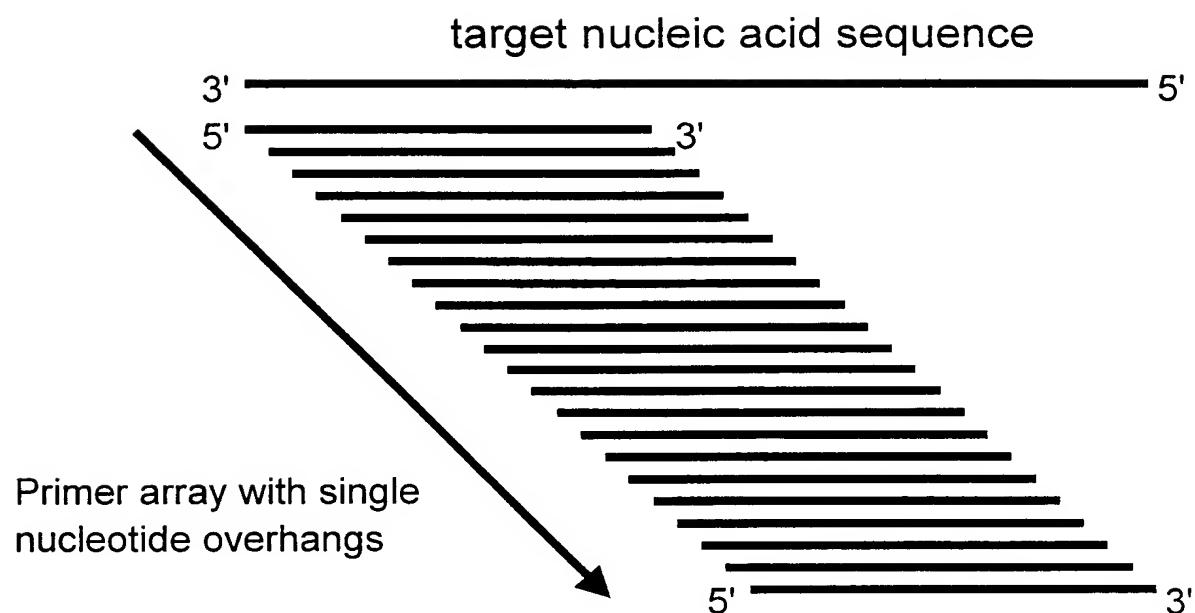
## CLAIMS

What is claimed is:

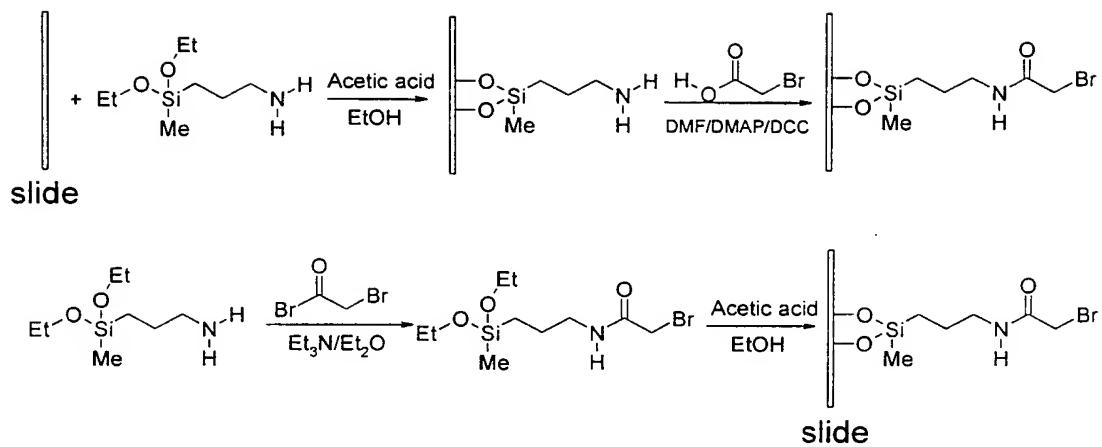
1. A method for detecting a mutation in a target nucleic acid sequence that comprises:
  - A. attaching oligonucleotide primers to a substrate, wherein the oligonucleotide primers have a sequence that is complementary to the target nucleic acid sequence, and wherein the oligonucleotide primers are grouped according to the identity of the first base which would be expected to be added to the primer through the process of primer extension;
  - B. hybridizing to the oligonucleotide primers a sample nucleic acid sequence which possibly contains a mutation;
  - C. extending each oligonucleotide primer by one base using a reaction mixture comprising labeled terminators and enzyme; and
  - D. detecting a mutation in the sample nucleic acid sequence by detecting the presence of a labeled terminator which does not correspond to the identity of the base expected to be added to the primer through the process of primer extension.
2. The method of Claim 1, wherein the sample nucleic acid sequence is mRNA.
3. The method of Claim 1, wherein the sample nucleic acid sequence is DNA.
4. The method of Claim 1, wherein the labeled terminators are selected from the group consisting of deoxynucleoside triphosphates and dideoxynucleoside triphosphates.
5. The method of Claim 4, wherein the labeled terminators are synthetically modified.
6. The method of Claim 1, wherein the labeled terminators are labeled with a fluorescent dye.
7. The method of Claim 1, wherein the labeled terminators are labeled with a chemiluminescent reagent.

8. The method of Claim 1, wherein the labeled terminators are labeled with a radioactive label.
9. The method of Claim 1, wherein the labeled terminators are labeled with a redox tag.
10. The method of Claim 1, wherein the labeled terminators are labeled with an electrically conductive tag.
11. A nucleic acid array comprising a substrate containing oligonucleotide primers having a sequence complementary to a target nucleic acid sequence, wherein the oligonucleotide primers are grouped according to the identity of the first base which would be expected to be added to the primer through the process of primer extension after hybridizing with a sample nucleic acid sequence.
12. A kit for detecting genetic mutations comprising the nucleic acid array according to Claim 11.
13. A method for detecting a mutation in a target nucleic acid sequence that comprises:
  - A. identifying the base expected to be added to a primer located at a particular coordinate on an oligonucleotide array as expected from the target nucleic acid sequence;
  - B. identifying the base actually added to the primer located at the particular coordinate on the oligonucleotide array through the process of primer extension;
  - C. comparing the base actually added to the primer at the particular coordinate with the base expected to be added to the primer at the particular coordinate; and
  - D. reporting those instances where the bases are not the same, in order to identify a mutation.

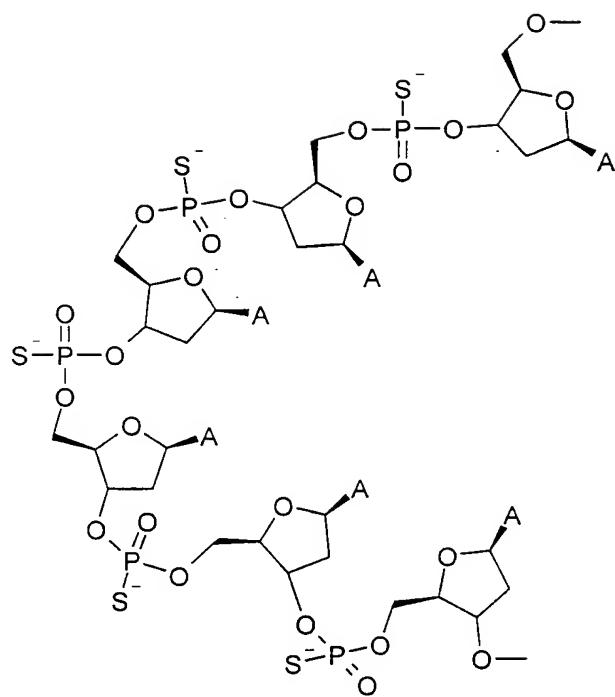
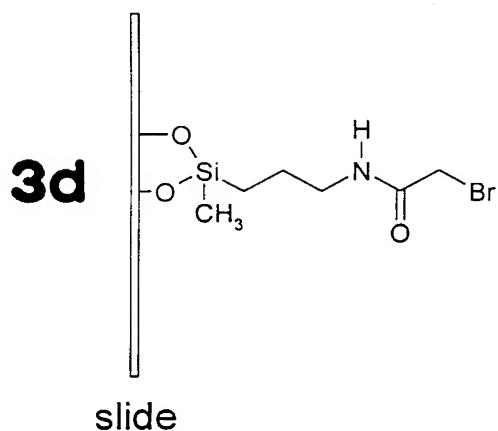
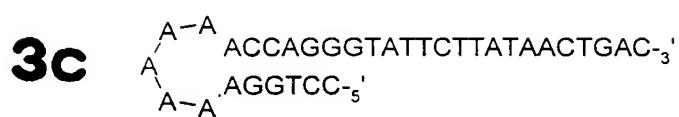
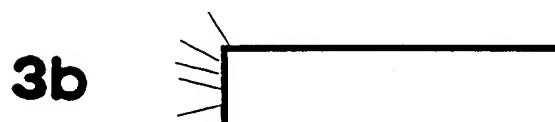
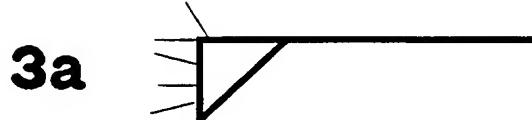
1/6



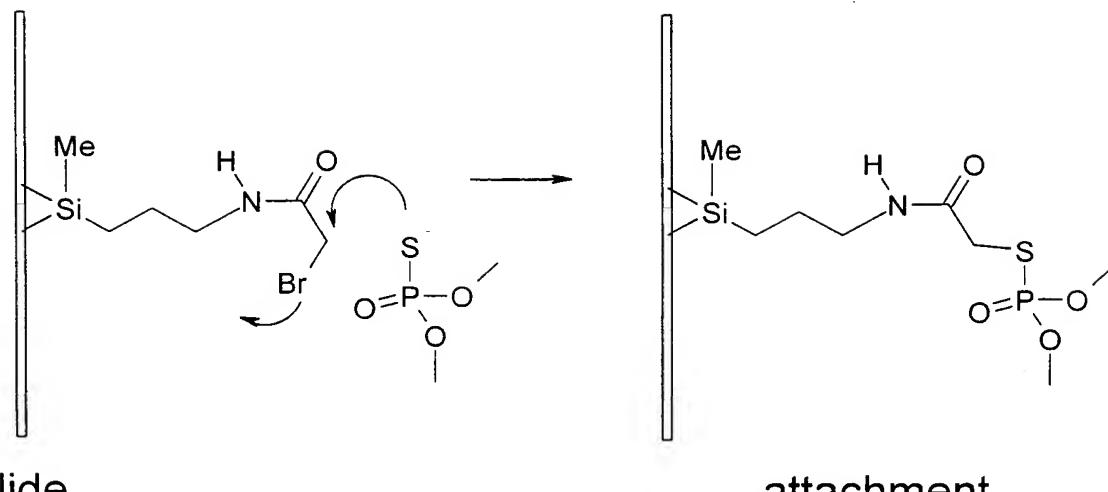
2/6



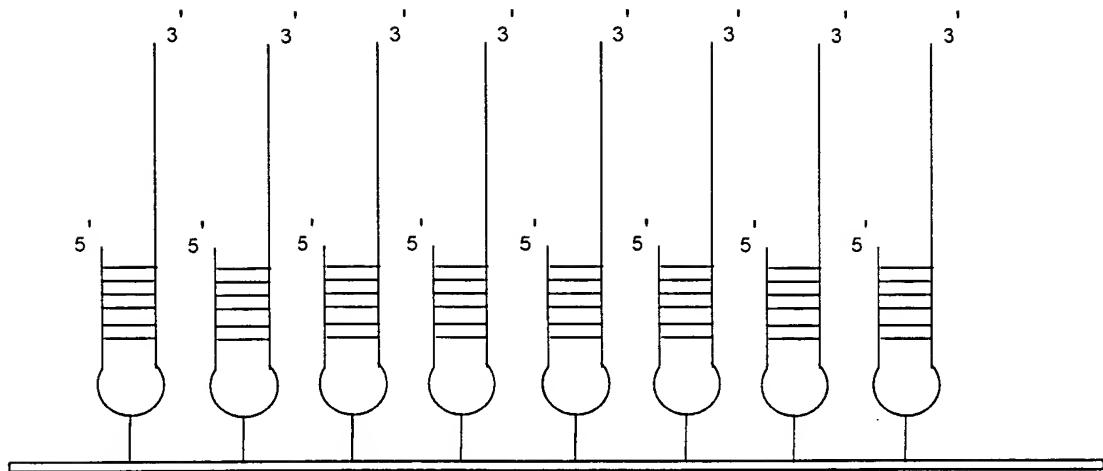
3/6



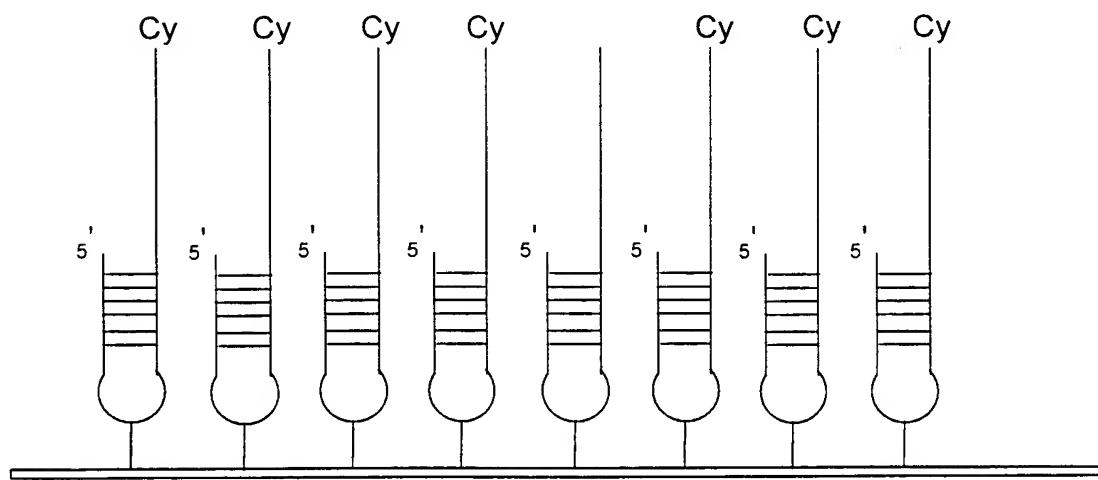
4/6



5/6



primer extension  
cy labeled ddNTPs



6/6

